

RESULTS AND REMARKS ON ANALYZING BIOLOGICAL REPEATED MEASURES

Michael P. Meredith¹

BU-956-M²

January 1988

ABSTRACT

The distribution of errors from a biological repeated measures (BRM) experiment often has a covariance matrix of Toeplitz form. For example, the covariance matrix for a stationary first-order autoregressive process is an element in the class of Toeplitz matrices. It is shown that univariate analyses of certain linear combinations of intra-subject repeated measures are pairwise uncorrelated provided the Toeplitz error assumption is valid. This class of linear combinations includes orthogonal polynomial contrasts as a special case. Additionally, a liberal lack-of-fit test is suggested. These results are useful for the successful conduct of a precise and easily interpreted analysis of BRM data. Usual recommendations for analysis of BRM experiments are briefly reviewed. Pitfalls inherent in either the assumptions or interpretations of the usual methods will be highlighted. These pitfalls include untenable assumptions and the inability to provide interpretations that directly correspond to the aims of the researcher.

Keywords: autoregressive errors, growth curve analysis, serial correlation, Toeplitz matrix.

¹ Assistant Professor of Biological Statistics, Biometrics Unit, Cornell University

² In the Biometrics Unit technical series, Cornell University, Ithaca, NY 14853

1. INTRODUCTION

1.1 Objectives

The purpose of this paper is twofold. First, three methods commonly utilized to analyze repeated measures from biological studies are briefly reviewed. Second, a helpful fact is reported about the analysis of a certain class of linear combinations of the repeated measures for each experimental unit (EU). This class of linear combinations includes orthogonal polynomial contrasts as a special case. It is seen that pairs of univariate analyses of these linear combinations are uncorrelated under an assumed Toeplitz error structure. This error structure includes the stationary first-order autoregressive process as a special case.

The three methods considered can be grouped into MANOVA, split-unit univariate analyses, and the analysis of coefficients of functions estimated from the t repeated measures on each EU. Some of the advantages and disadvantages or deficiencies are indicated for each of these methods.

At least part of the confusion in the analysis of repeated measurements in biological settings originates from an attempt to utilize methods intended for use on psychometrical or sociological data. The objectives of these analyses typically differ fundamentally from objectives of a biological repeated measures experiment. This confusion has been exacerbated by the influence of large statistical software packages on the user community. That is, people tend to use methods that can be accomplished easily on their mainframe, mini-, or micro-computer. These methods happen to be the split-plot univariate approach to repeated measures (Winer, 1971) and MANOVA profile analysis (Morrison, 1976). Both approaches fail to exploit the fact that the underlying response functions are continuous in time.

In biological repeated measures studies the interest generally focuses upon the nature and behavior of the underlying response functions. These response functions may be growth curves in feeding trials, curves describing the level of hormone in the blood of an animal during lactation, the behavior of enzyme kinetics (*in vivo* or *in vitro*), and so forth. The success of an analysis of these response functions depends upon replacing the repeated measures on each EU by a few estimated coefficients which result in the most effective and informative comparisons amongst treatments. It will be concluded that for many biological repeated measures experiments the simple univariate analysis of coefficients (often linear, but possibly nonlinear combinations of the responses) is the most useful

avenue to take. It is most useful because analyzing the coefficients of functions of basic interest results in inferences and interpretations that directly address the original questions of the research. Such results are not always immediately forthcoming from the multivariate approach or a Geisser-Greenhouse-Box (i.e., split-plot analysis) approach to repeated measures.

1.2 Data and Model

In this discussion it is assumed that a treatments are allocated to experimental units (EUs) in a completely randomized design (CRD). However, additional design features such as blocking (stratification) or covariates can be easily incorporated into some of the analyses. Without loss of generality it will be assumed that the a treatments are equally replicated with n EUs per treatment. Measurements are taken at t subsequent times for each of the EUs. Thus, there are ant observations under study. The statistical layout of the data is displayed in figure 1.1.

Further, it is assumed that the responses are measured on a scale, or have been successfully transformed to a scale, that has additive homoscedastic (normally distributed) errors. Note that the term "repeated measures" will *not* be applied to data arising from experiments where treatments are changed at subsequent time periods as in crossover trials (see comments by Yates, 1982).

Some notation and model assumptions are detailed below:

$$Y_{ant \times 1} = \begin{pmatrix} y_{11} \\ y_{21} \\ \vdots \\ y_{at} \end{pmatrix} \sim MVN_{ant} \left[\mu, I_{an} \otimes \Sigma_{t \times t} \right], \text{ each } y_{ik} \text{ is the } t \times 1 \text{ column vector}$$

of repeated measures on the ij^{th} experimental unit. A typical formulation would be $y_{ijk} = \mu_{ik} + \varepsilon_{ijk}$, for $i=1, 2, \dots, a$, $j=1, 2, \dots, n$, and $k=1, 2, \dots, t$, and where

$$\text{Cov}(y_{ijk}, y_{i'j'k'}) = \begin{cases} \sigma^2 & \text{for } i=i', j=j', \text{ and } k=k' \\ \rho_{kk'} \sigma^2 & \text{for } i=i', j=j', \text{ and } k \neq k' \\ 0 & \text{otherwise.} \end{cases}$$

Often it is helpful to assume some functional form for the μ_{ik} 's such as

$\mu_{ik} = f(\beta_i, k)$, for example a polynomial of degree $\leq t-1$.

In the above formulation the β_i are unknown parameter vectors that may result in substantial simplification of the μ_{ik} .

Figure 1.1: Statistical layout of repeated measures data discussed in this paper.

| Treatment | Exp'tl Unit | <u>Time</u> | | | |
|-----------|----------------|-------------|-----------|-----|-----------|
| | | 1 | 2 | ... | t |
| 1 | 1 | | | | |
| | 2 | | | . | . |
| | : | Y_{1j1} | Y_{1j2} | . | Y_{1jt} |
| | : | | | . | . |
| | n_1 | | | . | . |
| : | : | : | : | | : |
| : | : | : | : | | : |
| a | 1 | | | | |
| | 2 | | | . | . |
| | : | | | . | . |
| | : | Y_{aj1} | Y_{aj2} | . | Y_{ajt} |
| | n_a | | | . | . |

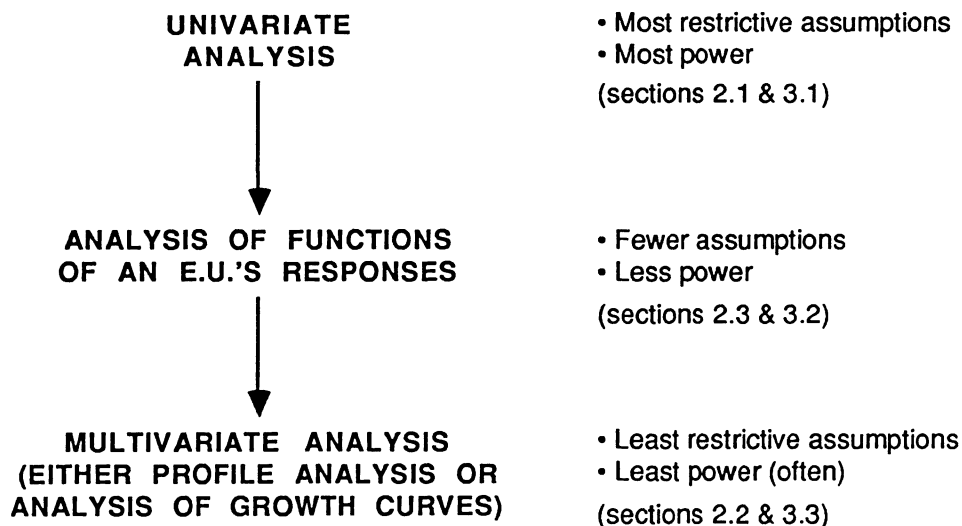
2. USUAL METHODS OF ANALYSIS

Three general methods of analysis are most commonly employed to analyze data from a repeated measures experiment. These methods will be broadly grouped into: 1) the multivariate analysis of variance; 2) split-unit type univariate analyses; and, 3) the analysis of coefficients from functions estimated from the t repeated measures on each EU.

There are, of course, numerous other suggestions for analyzing data from repeated measures studies. Many of these methods represent variations on the three themes indicated above. The naïve practice of performing a separate analysis of variance for the data at each time will not be considered. Further, if the number of repeated measures on each EU is quite large (say, $t \geq 25$) then it is likely the analysis will focus upon the properties of the underlying multivariate time series. This situation is not the focus of this paper and the reader is referred to Brillinger (1981). It is tempting under the assumption of a stationary first-order autoregressive error process to estimate ρ (the autocorrelation coefficient) and use estimated generalized least squares (EGLS) in the analysis. However, for a fixed number of measurements, t , the usual estimators of ρ are not consistent (even if the number of EUs is large). This methodology is used with success in econometric applications where t may be very large (see Parks, 1967).

In this section each of these methods will be described briefly. The ordering of the three potential approaches to the analysis of repeated measures is depicted in figure 2.1. In section 3 some of the advantages and disadvantages for each approach will be discussed.

Figure 2.1: Relationship between the three usual methods of analysis for repeated measures experiments.



2.1 Univariate Split-Unit Analysis

On one extreme the t repeated measures for each EU are treated as t sub-unit

measurements on each whole-unit (viz., EU). The subsequent univariate split-unit type of analysis requires not only the assumption of normality, but also that the errors are spherically symmetric (Huynh and Feldt, 1970). It may be shown that such an assumption is, for all practical purposes, equivalent to an assumption of equicorrelation (also termed compound symmetry by some authors). This assumption is generally untenable for analysis of repeated measures due to the presence of serial correlation. However, the univariate analysis has the advantage of being easily performed computationally as well as being easily interpreted after an appropriate partitioning of the sums of squares in the ANOVA table. Since spherically symmetric errors is usually an unrealistic assumption there has been considerable effort devoted to refining the conservative Greenhouse-Geisser-Box approach (see section 3.1). Many papers appear on this topic in the psychometrical literature and elsewhere (see Greenhouse and Geisser, 1959; Monlezun, Blouin, and Malone, 1984; Rogan, Keselman, and Mendoza, 1979; Huynh and Feldt, 1970; Huynh, 1976; Gill and Hafs, 1971; Milliken and Johnson, 1985; Frane, 1982; Elashoff, 1986).

2.2 Multivariate Analysis

The t repeated measures on each EU may be correctly regarded as a t -variate response vector and analyzed via methods of multivariate analysis. This is certainly a correct approach to the problem provided that t does not exceed the number of EUs for each treatment. Only assumptions of normality and equal variance-covariance matrices across the a treatment populations are required for valid hypothesis testing. No assumptions are made about the structure of the common variance-covariance matrix.

Profile analysis (Morrison, 1976; Davidson, 1980; Elashoff, 1986) and the multivariate analysis of growth curves (Rao, 1958, 1965; Potthoff and Roy, 1964; Khatri, 1966; Grizzle and Allen, 1969) are two of the most frequently employed approaches for looking at repeated measures data with multivariate techniques. Other pertinent papers include Cole and Grizzle (1966), and Danford, Hughes, and McNee (1960).

2.3 Analysis of Estimated Coefficients

Between the two extremes of the univariate split-unit analysis and the general multivariate analysis are intermediate methods that focus upon analyzing coefficients of functions fitted to the t repeated measures for each EU. These func-

tions may be simple linear combinations that are meaningful to the investigator, polynomials that serve to approximate the true unknown nonlinear response function within a specific time interval, or nonlinear models describing growth or the kinetics in pharmacological models. The analysis of linear contrasts has been discussed in many places (see Wishart, 1938; Box, 1950; Elashoff, 1986; Snedecor and Cochran, 1980; Rowell and Walters, 1976; Walters and Rowell, 1982; Yates, 1982; and, Zerbe, 1979).

It is important to note that *each* EU results in a set of estimated coefficients that are then treated as data. This set of coefficients may result in a reduction in the dimensionality of the response vectors which serves as an initial simplification to the problem. Furthermore, these functions are chosen by the investigator because they are of interest for a proper analysis of the data. Consequently, analyzing these coefficients will provide *direct* evidence for or against hypotheses set forth initially in the experiment. The analysis of estimated coefficients easily allows for restrictions on randomization (design features) as well as the incorporation of covariates in the analysis.

3. ADVANTAGES AND DISADVANTAGES

3.1 Univariate Split-Unit Analysis

The usual advantages cited for the split-unit univariate approach are that it is easily computed, easily understood, and easily interpreted. The wide availability of powerful statistical software tends to minimize the importance of the first reason. The second and third reasons are true enough, but they fail to consider whether or not the questions being addressed correspond to those of the investigator. Most often, the tests for time and time×treatment effects are of little or no interest to the researchers—they know that the response curves differ through time for the different treatments and interest generally focuses on the nature of the response curves. This information can often be obtained by further partitioning of the time and time×treatment sums of squares, but this requires the existence of a valid error term for testing the sub-hypotheses. Unfortunately, such a valid error term only exists under the rather unrealistic assumption of a spherically symmetric error structure.

The reason that a split-unit design (or more simply a randomized complete block design) has a valid error term is because each contrast in a complete orthonormal set of contrasts amongst sub-units has a common variance (viz., $[1-(t-1)\rho]\sigma^2$ under equicorrelation assumption) if there are t sub-units per EU (or

block). Thus, the variance estimates may be combined to form a pooled estimate of error. In the absence of randomization (e.g., if the sub-unit is time) this pooling is seen to be valid only if the assumption of spherically symmetric errors is true (see Huynh and Feldt, 1970).

Under the sphericity assumption, (i.e., for some orthonormal matrix C the condition $C\Sigma C' = \lambda I$ is true, where Σ is the variance-covariance matrix and λ is a scalar), the linear contrasts amongst times estimate a common error, even though it may contain an unknown function of V where $\Sigma = \sigma^2 V$. However, under more general error assumptions (e.g., the Toeplitz assumption to be discussed in section 4) different linear contrasts have different variances involving different polynomial functions of the unknown elements of V . Consequently, a pooled estimate cannot be determined in this case. For example, consider the following simple case where $t=3$.

$$\text{Let } \lambda_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \lambda_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \text{ and } \lambda_2 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix},$$

$$\text{also, let } \Sigma_1 = \begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{pmatrix} \sigma^2 \text{ and } \Sigma_2 = \begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{pmatrix} \sigma^2. \text{ Then we have the following:}$$

$$\text{Var}(\lambda_1' y) = \text{Var}(\lambda_2' y) = 2(1-\rho)\sigma^2 \text{ under assumption } \Sigma_1, \text{ whereas}$$

$$\text{Var}(\lambda_1' y) = 2(1-\rho)(1+\rho)\sigma^2 \neq \text{Var}(\lambda_2' y) = 2(1-\rho)(3-\rho)\sigma^2 \text{ under assumption } \Sigma_2.$$

The following example shows why equicorrelation yields meaningful estimates (details may be found in Arnold, 1981).

$$Y_{\text{ant} \times 1} \sim \text{MVN} \left\{ \mu, \sigma^2 V = \left(I_{\text{an}} \otimes [(1-\rho)I_{t \times t} + \rho J_t] \right) \sigma^2 \right\}$$

Consider the transformation $\Gamma_{\text{ant} \times \text{ant}} : Y_{\text{ant} \times 1} \rightarrow Y_{\text{ant} \times 1}^*$, where

$$\Gamma = \begin{bmatrix} \mathbf{I}_{an} \otimes \frac{1}{\sqrt{t}} \mathbf{j}'_t \\ \mathbf{I}_{an} \otimes \mathbf{C}'_{t-1 \times t} \end{bmatrix}, \text{ where } \mathbf{C}' \mathbf{C} = \mathbf{I}_{t-1}, \mathbf{C} \mathbf{C}' = \mathbf{I}_t - \frac{1}{t} \mathbf{J}_t, \text{ and}$$

$$\mathbf{C}' = (\mathbf{c}'_i) = \left(\frac{1}{\sqrt{i(i-1)}} \mathbf{j}'_{i-1}, \frac{-(i-1)}{\sqrt{i(i-1)}}, 0 \mathbf{j}'_{t-i} \right), \text{ the } i^{\text{th}} \text{ row of a Helmert matrix.}$$

$$\text{Then, } \mathbf{Y}_{ant \times 1}^* = \begin{bmatrix} \mathbf{Y}_1^* \\ \mathbf{Y}_2^* \end{bmatrix} \sim \text{MVN} \left\{ \boldsymbol{\mu}^*, \sigma^2 \begin{bmatrix} (1+(t-1)\rho) \mathbf{I}_{an} \\ (1-\rho) \mathbf{I}_{an(t-1)} \end{bmatrix} \right\}.$$

It is well known that positive serial correlation, the most common departure from the assumption of spherically symmetric (or more realistically equicorrelated) errors, results in overly liberal tests of the hypotheses associated with time and time \times treatment (see Hearne, Clark, and Hatch, 1983; Huynh and Feldt, 1976). Wallenstein and Fleiss (1979) show that under an assumption of a stationary first-order autocorrelated error process that the lower bound on the degrees of freedom associated with the time and time \times treatment F-tests can be increased from the bounds of $[1, a(n-1)]$ and $[a-1, a(n-1)]$, respectively, given by Box (1954b).

The spherically symmetric error assumption is not a particularly general condition when considering plausible error structures that would arise from biological repeated measures studies. For all practical purposes the assumption is truly one of equicorrelation (sometimes termed compound symmetry) as indicated by Yates (1982). The assumption of equicorrelation is clearly untenable for the overwhelming majority of biological repeated measures situations.

The spherically symmetric error assumption is particularly critical for univariate analysis of repeated measures for another reason. The validity of the within EUs error term is questionable due to the fact that time is a systematic factor that cannot be randomized. Thus, the usual argument of randomization tests that are given for blocked experiments does not apply to the repeated measures situation (Cox, 1958, and Scheffé, 1959).

The typical remedy for violations of the necessary error assumptions is to construct either a conservative F test as indicated above, or estimate the approximate degrees of freedom {as $[(t-1)\epsilon, a(n-1)(t-1)\epsilon]$ and $[(a-1)(t-1)\epsilon, a(n-1)(t-1)\epsilon]$,

where $(t-1)^{-1} \leq \epsilon \leq 1$ depending upon an estimate of the measure of departure from sphericity, ϵ (see Huynh and Feldt, 1976). Of course, estimation of ϵ will introduce even further uncertainty into the problem. Many Monte Carlo studies have been conducted to examine the behavior of such estimation. The frequently advocated conservative F-tests approach are predicated on the papers of Box (1954a, 1954b). These conservative tests and the tests which estimate ϵ are currently standard fare in the output of most major statistical packages.

The relative power of the univariate split-unit approach depends upon how many repeated measures are taken (i.e., the size of t). If t is small (e.g., $t=2$ or 3) then there is little or nothing to be gained in terms of power by using the univariate split-unit approach. For example, when $t=2$ the proper analysis consists of simply examining the within EU sums and differences. It must be noted that the degrees of freedom associated with the conservative tests correspond exactly to those found in the univariate analysis of estimated coefficients. The important distinction being that the analysis of coefficients focuses on questions of direct interest to the investigator. If the conservative test of time \times treatment effects is rejected in the split-unit approach then the investigator is at a loss as to how to proceed further. There are no valid estimates of standard errors to attach to simple effect comparisons. This is a serious limitation of the split-unit approach when the sphericity assumption is not met.

A common strategy for the split-unit approach is to conduct preliminary tests to ascertain the validity of the sphericity and homogeneity assumptions as indicated below in the null hypotheses H_{01} and H_{02} . Such a strategy results in serious bias of the significance level in the final test that is made. These conditional specification problems have been largely ignored in the repeated measures literature (see Bancroft and Han, 1977, for a review and bibliography of conditional specification).

The following example shows what a critical function, $\phi(Y)$, may look like as a composition of the critical functions for all the conditional tests involved. Clearly, the level of the resulting test is unknown.

$$H_{01}: C\Sigma_i C' = C\Sigma_{i'} C', \text{ for all } i \neq i'.$$

$$H_{02}: C\Sigma C' = \lambda I$$

$$H_{03}: \text{no time} \times \text{treatment interactions}$$

$$H_{04}: \text{parallel profiles}$$

$$\phi_1(Y) = \begin{cases} 1 & \text{if } H_{01} \text{ is rejected,} \\ 0 & \text{otherwise.} \end{cases} \quad \phi_2(Y) = \begin{cases} 1 & \text{if } H_{02} \text{ is rejected,} \\ 0 & \text{otherwise.} \end{cases}$$

$$\phi_3(Y) = \begin{cases} 1 & \text{if } H_{03} \text{ is rejected,} \\ 0 & \text{otherwise.} \end{cases} \quad \phi_4(Y) = \begin{cases} 1 & \text{if } H_{04} \text{ is rejected,} \\ 0 & \text{otherwise.} \end{cases}$$

$$\phi(Y) = [1 - \phi_1(Y)][1 - \phi_2(Y)]\phi_3(Y) + [1 - \phi_1(Y)]\phi_2(Y)\phi_4(Y)$$

Furthermore, there is typically very low power for testing the validity of the sphericity assumption using Mauchley's (1940) criterion (or Grieve, 1984). However, the level of the F test for testing time and time \times treatment effects is very sensitive to extremely small departures from the sphericity assumption (see Keselman, Rogan, Mendoza, and Breen, 1980; Boik, 1981). To avoid such problems Boik (1981, p. 254) recommends that "A reasonable alternative is to employ a separate estimate of experimental error for each contrast among the repeated measures. ...The loss in degrees of freedom [from $a(n-1)(t-1)$ to $a(n-1)$] is more than compensated for by the gain in control over test size and power." This is *precisely* the idea of analyzing meaningful estimated coefficients from functions (e.g., contrasts) fitted to each EU.

3.2 Analysis of Estimated Coefficients

The primary advantage of the analysis of estimated coefficients is that questions of interest, formulated *by the investigator* in terms of response functions or other contrasts, are addressed directly by this approach. One very important result of this analysis is that coefficients and their standard errors are calculated. (Remember that the conservative split-plot approach enabled *only* hypothesis testing.) These may then be used to display the response curves for the treatments under study. For the majority of biological repeated measures experiments this avenue would seem to be the most favorable to elicit the proper information from the data collected. It is easily understood and easily undertaken. However, this approach cannot be pre-programmed into any statistical software package because the response functions (or any other combination of interest, linear or nonlinear) are necessarily specified by the investigator to address the needs of the particular process under study. If these response functions are reflected in orthogonal polynomial contrasts then several

statistical software packages do provide this decomposition (sums of squares only, no estimates) as an option.

One possible problem with this approach is that few authors seem to emphasize that the univariate analyses based upon a set of orthogonal contrasts are generally *not* independent of one another. That is, even though the coefficients for two linear combinations, say λ_1 and λ_2 , are orthogonal, this does not guarantee independent analyses. Consequently $\lambda_1'\lambda_2=0$ does not imply $\lambda_1'\Sigma\lambda_2=0$ unless the sphericity assumption is valid. For example, it is easy to demonstrate that the quadratic forms for the sum and quadratic orthogonal polynomial contrast are not independent for a first-order autoregressive error process with $t=3$ equally spaced repeated measures. This simple example is demonstrated below.

$$\text{Let } \Sigma = \begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{pmatrix} \sigma^2, \text{ and let } \lambda_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \text{ and } \lambda_2 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \text{ then } \lambda_0'\Sigma\lambda_2 \text{ is}$$

$$(1 \ 1 \ 1) \begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = (1+\rho+\rho^2 \ 1+2\rho \ 1+\rho+\rho^2) \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = 2\rho(\rho-2) \neq 0$$

The following section will describe some independence results for a useful set of linear combinations amongst repeated measures that have a fairly general serial correlation error structure. If the estimated coefficients to be analyzed are not independent then it is wise to either proceed with a multivariate analysis of the coefficients (easier to conduct and interpret than the usual MANOVA approach), or conduct separate univariate analyses of the coefficients and adjust the error levels using Bonferroni bounds (Feller, 1968).

Another advantage of this approach is that additional design features and covariates can be included easily in the analysis. For example, if the response being analyzed is $\lambda_1'y$ then a covariate x that was measured at each time period would be included as $\lambda_1'x$ in the analysis. Likewise, blocking or stratification present no difficulty for the analysis.

One additional feature should be noted. For linear estimators the coefficient estimates are generally inefficient but unbiased for the coefficients of interest. This is because the error structure has (usually) not been taken into account when estimating the coefficients. The lack of efficiency may be ameliorated in one of several ways. First, a well informed guess at the underlying error structure would result in more efficient (and still unbiased) estimators for each EU. The usual approach is to simply increase the total number of subjects to improve precision. A combination of these two is the most rational path to follow.

In the case of nonlinear combinations there may also be some bias in the estimated coefficients. This is no surprise as nonlinear least squares estimation typically results in biased, though consistent, estimators. For nonlinear functions the estimated expected values of the coefficients will be biased due to the fact that $E[g(Y)] \neq g(E[Y])$ for nonlinear functions $g(\cdot)$. Jensen's inequality may be used to gain some information about the bias depending upon the convexity of the function $g(\cdot)$. This is generally not perceived to be a serious problem in practice although it should be considered for each individual problem if nonlinear functions are being utilized and they cannot be linearized by some appropriate transformation.

3.3 Multivariate Analyses

The t responses through time for each EU are correctly regarded as a t -variate response vector. Thus, the multivariate analysis of variance is a natural way to proceed when analyzing repeated measures data. The greatest advantage is that no assumptions need be made about the error structure of the repeated measures other than equal variance-covariance matrices across the normally distributed treatment populations. However, this advantage also results in one of the methods greatest disadvantages—low power due to having to estimate the underlying variance-covariance matrix Σ with S , the sample covariance matrix. This is an extremely important consideration if t is large and n is not.

The above problem is disconcerting because realistic assumptions can usually be made for the properties of the errors from biological repeated measures experiments. For example, some sort of positive serial correlation is most likely to be present. The problem may be resolved by estimating Σ under the assumption of some patterned variance-covariance structure such as serial correlation or a Toeplitz pattern. This constrained estimation would result in higher power of the resulting tests and confidence ellipsoids with smaller volumes. In practice this constrained

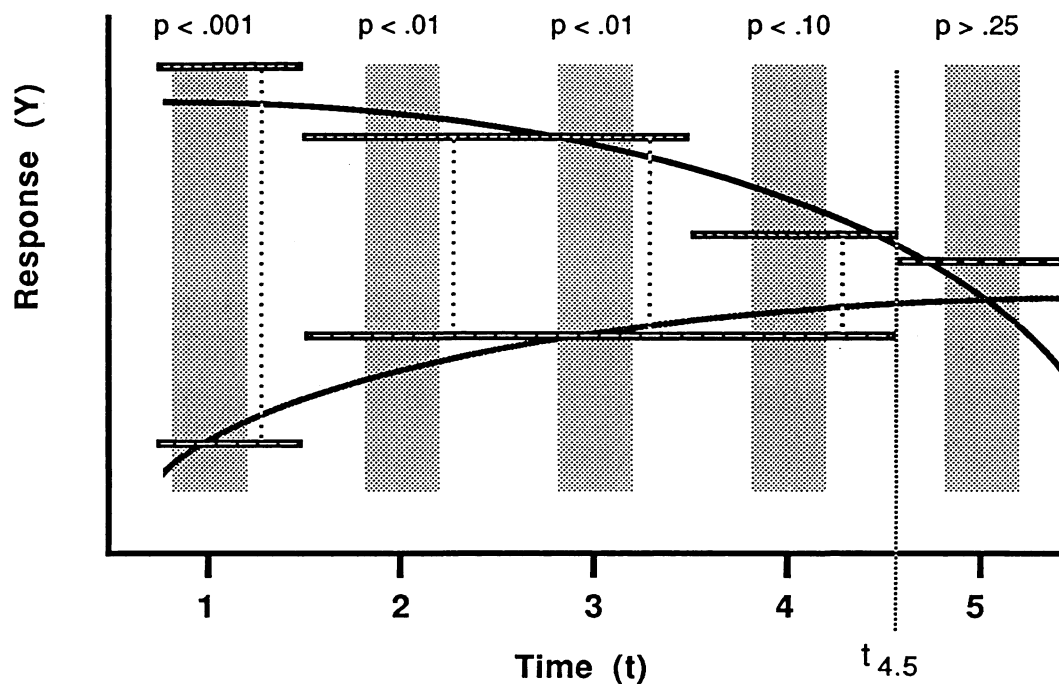
estimation is not always easily accomplished due to difficulties in maximizing the necessary maximum likelihood equations. It has been shown in section 3.2 how to capitalize on some patterned variance-covariance structures when analyzing coefficients.

Another problem is that most texts and statistical software focus on profile analysis. An emphasis on profile analysis leads an investigator to examine the hypotheses of interest indirectly rather than taking an objective approach. The user should proceed with an analysis that directly addresses the questions of research interest. Further, the hypothesis of no time \times treatment effects is known to be false in almost all non-trivial biological contexts. Thus, the analysis must be directed to the nature of these "interaction" effects as reflected in differences amongst coefficients of meaningful response functions. However, Morrison (1976, p. 208) recommends that, "If the hypothesis of parallel-treatment population profiles cannot be accepted, it will be necessary to test the equality of the treatment effects separately for each response by t univariate analyses of variance." This advice is counter-productive for making useful inferences in typical biological situations.

The foremost criticism of the above recommendation is that the time at which significant differences are detected is simply a function of the number of replications of EUs. Change the number of replications sufficiently and you can change the point in time at which "significance" is detected. Gill (1986) makes a similarly dubious recommendation for the univariate split-plot approach.

The following figure 3.1 attempts to illustrate the fallacy behind performing analyses at each time period. The solid black continuous curves represent the true underlying response functions. The step functions indicate thresholds between time periods as suggested by the p -values associated with t -tests performed at each time point. One threshold, labelled $t_{4.5}$, is shown.

Figure 3.1: Threshold versus continuous response functions (see text).



The MANOVA approach can be quite helpful when used to analyze the estimated coefficients as indicated in section 3.2. The multivariate analysis at this level generally retains good power since the number of coefficients that make up the multivariate response is usually much less than t . A good example of this kind of analysis on data from dairy cows is discussed in Allen, Burton, and Holt (1983), and Allen (1983).

4. CONTRASTS THAT GIVE UNCORRELATED ANALYSES

This section gives a result that is often helpful in the analysis of repeated measures experiments. A general result for pairs of orthogonal linear combinations is given that results in the univariate analyses of these pairs being independent when the underlying error structure is that of a Toeplitz form as defined below.

Definition: A real symmetric matrix V is said to be a real Toeplitz matrix if the elements of V , ρ_{ij} , satisfy the following condition

$$\rho_{ii'} = \begin{cases} \rho_{|i-i'|} & \text{for } i \neq i', \\ 1 & \text{for } i = i'. \end{cases}$$

As an example consider the variance-covariance matrix $\sigma^2 \mathbf{V}$ of a stationary first-order autoregressive process. The matrix \mathbf{V} has elements $\rho_{ij} = \rho^{|i-j|}$, and this process is seen to be of the Toeplitz form defined above. More general stationary serially correlated processes are included under this Toeplitz form.

Theorem: Let \mathbf{y} be a $t \times 1$ random vector from a multivariate normal (MVN) distribution with mean μ and variance-covariance matrix $\Sigma = \sigma^2 \mathbf{V}$ which satisfies the Toeplitz property. Let λ_1 and λ_2 be two $t \times 1$ vectors of known coefficients such that $\lambda_1' \lambda_2 = 0$, and λ_i ($i=1,2$) are of the form:

$$t \text{ odd: } \lambda_1 = \begin{bmatrix} c_1 \\ k \\ \mathbf{D} c_1 \end{bmatrix}, \text{ and } \lambda_2 = \begin{bmatrix} c_2 \\ 0 \\ -\mathbf{D} c_2 \end{bmatrix},$$

$$t \text{ even: } \lambda_1 = \begin{bmatrix} c_1 \\ \mathbf{D} c_1 \end{bmatrix}, \text{ and } \lambda_2 = \begin{bmatrix} c_2 \\ -\mathbf{D} c_2 \end{bmatrix}.$$

In the above, c_i ($i=1,2$) is a $[(t-1)/2] \times 1$ vector if t is odd, and c_i ($i=1,2$) is a $[t/2] \times 1$ vector if t is even, k is a known scalar. The matrix \mathbf{D} is a $[(t-1)/2]$ square matrix if t is odd, and a $[t/2]$ square matrix if t is even. The matrix \mathbf{D} is defined as follows:

$$\mathbf{D}_{q \times q} = [d_{ij}] = \begin{cases} 1 & \text{if } i+j-1=q \\ 0 & \text{otherwise.} \end{cases}$$

The matrix \mathbf{D} simply serves to reverse the ordering of the vector c_i ($i=1,2$). The conclusion is that the linear combinations, $\lambda_1' \mathbf{y}$ and $\lambda_2' \mathbf{y}$, are independent.

Proof: To prove the above fact it is sufficient to show that $\lambda_1' \mathbf{V} \lambda_2 = 0$, (Searle, 1971) for the specified forms of λ_i and \mathbf{V} a real Toeplitz matrix. The details of the proof

are omitted as they are straightforward though tedious. ■

A useful corollary to the preceding theorem is given by:

Corollary: Let \mathbf{y} be a $t \times 1$ random vector from a multivariate normal distribution with mean μ and covariance matrix $\Sigma = \sigma^2 \mathbf{V}$. If $\Sigma = \sigma^2 \mathbf{V}$ for \mathbf{V} a real Toeplitz matrix; and if λ_i , $i = 0, 1, 2, \dots, t-1$ are the orthogonal polynomial contrasts for equally spaced t , then $\lambda_i' \mathbf{y}$ and $\lambda_{i+2j+1}' \mathbf{y}$ are independent.

Proof: This follows immediately upon recognizing that orthogonal polynomials are of the required form in the above theorem. ■

Thus, under the assumptions of the theorem the even and odd powers of orthogonal polynomial contrasts for equally spaced time intervals satisfy the conditions for λ_1 and λ_2 . These particular contrasts prove to be useful in many equally spaced repeated measures studies. For example, the analyses of the means of EUs and the linear trends are uncorrelated, or the linear and quadratic components are uncorrelated. This is demonstrated for $t=3$ and a first-order autoregressive process as follows:

$$\text{Let } \Sigma = \begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{pmatrix} \sigma^2, \text{ and } \lambda_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \text{ and } \lambda_2 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \text{ then}$$

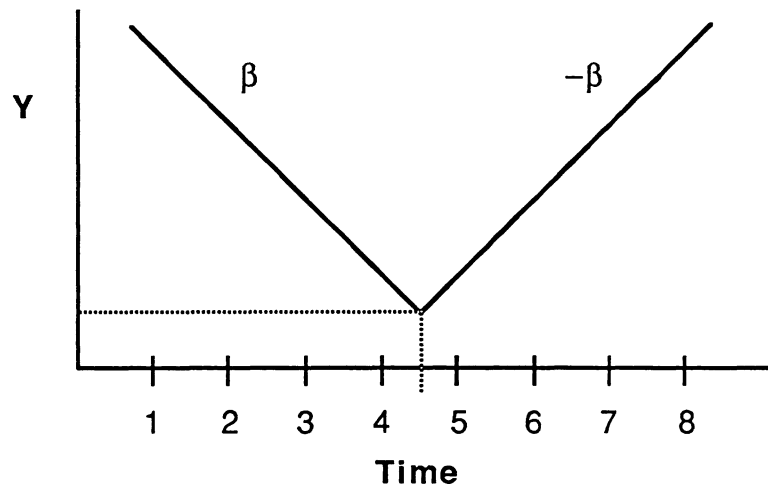
$$\frac{\lambda_1' \Sigma \lambda_2}{\sigma^2} = (-1 \ 0 \ 1) \begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = (-1 + \rho^2 \ 0 \ 1 - \rho^2) \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = 0.$$

Any linear combinations that satisfy the requirements of the theorem may be used and their analyses will be independent. This is a nice feature for making valid conclusions from repeated measures data since several simple contrasts may suffice to summarize the observed response functions. Often these contrasts are chosen to correspond to the linear trend and curvature.

In addition, one may partition the remaining sums of squares to assess lack-of-fit. This test would be performed using the error term associated with only the remaining lack-of-fit sums of squares. Thus, a liberal test for lack-of-fit would result assuming there exists some degree of positive serial correlation. Consequently, if there is no evidence for lack-of-fit from this test then systematic departures are of no detectable consequence.

The utility of the theorem is not limited to the class of orthogonal polynomial contrasts. Other linear combinations chosen to investigate segmented regressions, plateau response, and periodic response are within the class described in the theorem. For example, for $t=8$ equally spaced times let $\lambda'_1 = (3 \ 1 \ -1 \ -3 \ -3 \ -1 \ 1 \ 3)$ and $\lambda'_2 = (3 \ 1 \ -1 \ -3 \ 3 \ 1 \ -1 \ -3)$. Then λ'_1 and λ'_2 estimate the common slope β , and the difference of the slopes, respectively, for a segmented regression depicted in figure 4.1 below. Clearly, λ'_1 and λ'_2 satisfy the required form in the theorem.

Figure 4.1: A potential application of theorem, e.g., titration curve.



5. CONCLUSIONS

The analysis of repeated measures is one of the most active areas of research in applied statistics. Research and methodology papers abound in the mainstream statistical literature and in related literature from biometrics, psychometrics, epidemiology, and econometrics. This is not surprising since resources are limited in most fields and it is necessary to take repeated measures on the same experimental units rather than have independent groups at each time period. Additionally, the

idea of measuring the same experimental units through time has intuitive appeal to most investigators since the within EU variability should be smaller than the between EU variability.

The amount of effort devoted to devising sound methodology for analyzing repeated measures is certainly warranted. Consider the spectrum of possibilities that can arise in practice: measures not taken at the same time intervals for each EU, dichotomous or polychotomous response variables, truncated or censored data, missing observations resulting in unbalanced data, threshold response and change in functional form, correlation between concomitant variables and randomly observed unequal time intervals, and so forth.

This paper has investigated the specific case where EUs are measured at the same time intervals and the measurements are complete for each EU. Even in this restricted situation the range of possible analyses is intimidating to many that must conduct statistical analyses of repeated measures data. This problem is partly compounded by the methodologies available in, and thus inadvertently promoted by, major statistical software packages.

It was seen that for most biological experiments involving repeated measures the usual profile analysis approach is unsatisfactory for directly addressing the questions of real research interest. Further, the conservative F-test of the Greenhouse-Geisser-Box univariate approach tests hypotheses that are of little or no use for answering the questions of primary research interest. If one follows the usual advice and performs preliminary tests for sphericity and equal covariance matrices (both highly sensitive to the assumption of normality) then the resulting critical functions of the final hypotheses tested have unknown level of significance. In the MANOVA approach the user may not fare any better. If he or she performs a profile analysis then the same problems remain concerning the utility of the tested hypotheses. The only real change is that the sphericity assumptions are no longer necessary, although the tests may be considerably less sensitive than their univariate counterparts if the sphericity assumption is valid. The MANOVA analysis of growth curves (or any response functions linear in the parameters) is a valid approach for many biological repeated measures problems. The problems with such an approach stem mainly from the difficulties many users encounter in conducting a correct analysis and then providing a cogent, easily understood interpretation of the results.

Researchers in biological studies are most often interested in eliciting pertinent information about the nature of the response functions that underlie the process under scrutiny. The coefficients of these functions contain the information that can be directly understood and interpreted by the investigator for his or her problem. When the function chosen by the investigator has been fitted to the repeated measures for each EU then the data are reduced to a smaller collection of estimated coefficients. The univariate analyses of these coefficients are both easy to conduct and easy to interpret. However, these analyses are not generally independent of one another, even in the case where the functions represent orthogonal polynomial contrasts of the responses for each EU. Thus, it may be necessary to analyze these sets of coefficients as a multivariate response vector. It was shown that for a certain class of linear combinations that pairs of univariate analyses are uncorrelated when the underlying error structure has a variance-covariance matrix of the Toeplitz form. This class of linear combinations was seen to include orthogonal polynomial contrasts as a special case. This Toeplitz error structure includes the stationary first-order autoregressive process as a special case and is a reasonable approximation to the error structure for many biological repeated measures studies.

Implicit in the above recommendation is the *KISS* mandate for data analysis—Keep It Statistically Simple! The analysis of coefficients from meaningful response functions fitted to the repeated measures for each experimental unit satisfies this mandate for the statistical consumer while simultaneously providing a useful and powerful analysis of biological repeated measures.

REFERENCES

- Aitkin, M. (1981), "Regression Models for Repeated Measurements," *Biometrics*, 37, 831-832.
- Allen, O. B. (1983), "A Guide to the Analysis of Growth Curve Data with Special Reference to SAS," *Computers and Biomedical Research*, 16, 101-115.
- Allen, O. B., Burton, J.H., and Holt, J.D. (1983), "Analysis of Repeated Measurements from Animal Experiments Using Polynomial Regression," *Journal of Animal Science*, 57, 765-770.
- Andersen, A.H., Jensen, E.B., and Schou, G. (1981), "Two-Way Analysis of Variance with Correlated Errors," *International Statistical Review*, 49, 153-167.

- Arnold, S.F. (1981), *The Theory of Linear Models and Multivariate Analysis*, New York: John Wiley.
- Bancroft, T.A., and Han, Chien-Pae. (1977), "Inference Based on Conditional Specification: A Note and a Bibliography," *International Statistical Review*, 45, 117-127.
- Bock, R.D. (1975), *Multivariate Statistical Methods in Behavioral Research*, New York: McGraw-Hill.
- Boik, Robert J. (1981), "A Priori Tests in Repeated Measures Designs: Effects of Nonsphericity," *Psychometrika*, 46, 241-255.
- Box, G.E.P. (1950), "Problems in the Analysis of Growth and Wear Curves," *Biometrics*, 6, 362-389.
- Box, G.E.P. (1954a), "Some Theorems on Quadratic Forms in the Study of Analysis of Variance Problems, I. Effect of Inequality of Variance in the One-Way Classification," *Annals of Mathematical Statistics*, 25, 290-302.
- Box, G.E.P. (1954b), "Some Theorems on Quadratic Forms in the Study of Analysis of Variance Problems, II. Effect of Inequality of Variances and Correlation Between Errors in the Two-Way Classification," *Annals of Mathematical Statistics*, 25, 484-498.
- Brillinger, David R. (1981), *Time Series: Data Analysis and Theory* (revised edition), San Francisco: Holden-Day, Inc.
- Cole, J.W.L., and Grizzle, J.E. (1966), "Applications of Multivariate Analysis of Variance to Repeated Measurements Experiments," *Biometrics*, 22, 810-828.
- Cox, D.R. (1958), *Planning of Experiments*, New York: John Wiley.
- Danford, M.B., Hughes, H.M., and McNee, R.C. (1960), "On the Analysis of Repeated Measurements Experiments," *Biometrics*, 16, 547-565.
- Davidson, M. (1980), "The Multivariate Approach to Repeated Measures," Technical Report #75, BMDP Statistical Software, Los Angeles, CA.
- Elashoff, Janet D. (1986), "Analysis of Repeated Measures Designs," Technical Report #83, BMDP Statistical Software, Los Angeles, CA.
- Feller, W. (1968), *An Introduction to Probability Theory and Its Applications*, 3rd edition, volume 1, New York: John Wiley.

- Frane, J.W. (1980), "The Univariate Approach to Repeated Measures—Foundation, Advantages, and Caveats (preliminary version)," Technical Report #69, BMDP Statistical Software, Los Angeles, CA.
- Geisser, S., and Greenhouse, S.W. (1958), "An Extension of Box's Results on the Use of the F-Distribution in Multivariate Analysis," *Annals of Mathematical Statistics*, 29, 885-891.
- Gill, J.L. (1986), "Repeated Measurement: Sensitive Tests for Experiments with Few Animals," *Journal of Animal Science*, 63, 943-954.
- Gill, J.L., and Hafs, H.D. (1971), "Analysis of Repeated Measurements of Animals," *Journal of Animal Science*, 33, 331-336.
- Greenhouse, S.W., and Geisser, S. (1959), "On Methods in the Analysis of Profile Data," *Psychometrika*, 24, 95-112.
- Grieve, A.P. (1984), "Tests of Sphericity of Normal Distributions and the Analysis of Repeated Measures Designs," *Psychometrika*, 49, 257-267.
- Grizzle, James E., and Allen, David M. (1969), "Analysis of Growth and Dose Response Curves," *Biometrics*, 25, 357-381.
- Hearne, E.M., Clark, G.M., and Hatch, J.P. (1983), "A Test for Serial Correlation in Univariate Repeated Measures Analysis," *Biometrics*, 39, 237-243.
- Huynh, H., and Feldt, L.S. (1970), "Conditions Under Which Mean Square Ratios in Repeated Measurements Designs Have Exact F-Distributions," *JASA*, 65, 1582-1589.
- Huynh, H., and Feldt, L.S. (1976), "Estimation of the Box Correction for Degrees of Freedom From Sample Data in the Randomized Block and Split-plot Designs," *Journal of Educational Statistics*, 1, 69-82.
- Keselman, H.J., Rogan, J.C., Mendoza, J.L., and Breen, L.J. (1980), "Testing the Validity Conditions of Repeated Measures F Tests," *Psychological Bulletin*, 87, 479-481.
- Khatri, Chinubhai G. (1966), "A Note on the MANOVA Model Applied to Problems in Growth Curve," *Annals of the Institute of Statistical Mathematics*, 18, 75-86.
- Laird, N.M., and Ware, James H. (1982), "Random-Effects Models for Longitudinal Data," *Biometrics*, 38, 963-974.

- Mauchley, J.W. (1940), "Significance Test for Sphericity of a Normal n-Variate Distribution," *Annals of Mathematical Statistics*, 11, 204-209.
- Milliken, G. A., and Johnson, D. E. (1984), *Analysis of Messy Data* (Chapter 26), Belmont, CA: Lifetime Learning Publications.
- Monlezun, C.J., Blouin, D.C., and Malone, L.C. (1984), "Contrasting Split Plot and Repeated Measures Experiments and Analyses," *The American Statistician*, 38, 21-31.
- Morrison, Donald F. (1976), *Multivariate Statistical Methods* (2nd edition, Chapter 5), New York: McGraw-Hill.
- Olkin, I., and Vaeth, M. (1981), "Maximum Likelihood Estimation in a Two-Way Analysis of Variance with Correlated Errors in One Classification," *Biometrika*, 68, 653-660.
- Parks, R.W. (1967), "Efficient Estimation of a System of Regression Equations when Disturbances are Both Serially and Contemporaneously Correlated," *JASA*, 62, 500-509.
- Potthoff, Richard F., and Roy, S.N. (1964), "A Generalized Multivariate Analysis of Variance Model Useful Especially for Growth Curve Problems," *Biometrika*, 51, 313-326.
- Rao, C.R. (1958), "Some Statistical Methods for Comparison of Growth Curves," *Biometrics*, 14, 1-17.
- Rao, C.R. (1965), "Theory of Least Squares When the Parameters Are Stochastic and Its Application to the Analysis of Growth Curves," *Biometrika*, 52, 447-458.
- Rogan, J.C., Keselman, H.J., and Mendoza, J.L. (1979), "Analysis of Repeated Measurements," *British Journal of Mathematical and Statistical Psychology*, 32, 269-286.
- Rowell, J.G., and Walters, D.E. (1976), "Analysing Data with Repeated Observations on Each Experimental Unit," *Journal of Agricultural Science, Cambridge*, 87, 423-432.
- Searle, S.R. (1971), *Linear Models*, New York: John Wiley.
- Scheffé, H. (1959), *The Analysis of Variance*, New York: John Wiley.

Snedecor, G.W., and Cochran, W.G. (1980), *Statistical Methods* (7th ed., pp. 330-332), Ames, Iowa: The Iowa State University Press.

Wallenstein, S., and Fleiss, Joseph L. (1979), "Repeated Measurements Analysis of Variance When the Correlations Have a Certain Pattern," *Psychometrika*, 44, 229-233.

Wallenstein, S. (1982), "Regression Models for Repeated Measurements: Reader Reaction," *Biometrics*, 38, 849-850.

Walters, D.E., and Rowell, J.G. (1982), "Comments on a Paper by I. Olkin and M. Vaeth on Two-Way Analysis of Variance with Correlated Errors," *Biometrika*, 69, 664-666.

Ware, James H. (1985), "Linear Models for the Analysis of Longitudinal Studies," *The American Statistician*, 39, 95-101.

Winer, B.J. (1971), *Statistical Principles in Experimental Design* (2nd ed.), New York: McGraw-Hill.

Wishart, J. (1938), "Growth-Rate Determinations in Nutrition Studies with the Bacon Pig, and Their Analysis," *Biometrika*, 30, 16-28.

Yates, F. (1982), "Regression Models for Repeated Measurements: Reader Reaction," *Biometrics*, 38, 850-853.

Zerbe, Gary O. (1979), "Randomization Analysis of the Completely Randomized Design Extended to Growth and Response Curves," *JASA*, 74, 215-221.